

Jurnal Pendidikan Ekonomi & Akuntansi

Vol. 3, No. 2, Juni 2010

ISSN 1978-8770

1

Pendidikan Karakter Melalui Pembelajaran Ekonomi

Teguh C. Dalyono

9

Karakteristik Perangkat Tes Teori Kejuruan SMK Bisnis dan Manajemen Program Keahlian Akuntansi di Daerah Istimewa Yogyakarta

Natalina Premastuti Brataningrum

23

Persepsi Siswa Tentang Kebijakan Penyelenggaraan dan Proses Pembelajaran di Rintisan Sekolah Bertaraf Internasional (RSBI), Studi Kasus Siswa di SMA Berstatus RSBI di Wilayah Kota Yogyakarta, Kabupaten Sleman, dan Kabupaten Bantul Yogyakarta

B. Indah Nugraheni

39

Peningkatan Pemahaman Mahasiswa Tentang Materi Laporan Arus Kas dengan Menerapkan Metode *Problem-Based Learning*

Ignatius Bondan Suratno

51

Upaya Meningkatkan Pemahaman Siswa Pada Materi Pembelajaran Siklus Akuntansi Perusahaan Jasa Melalui Penerapan Metode Pembelajaran Kooperatif

L. Saptono

J. PEA	Vol. 3	No.2	Halaman 1 - 70	Yogyakarta Juni 2010	ISSN 1978-8770
--------	--------	------	-------------------	-------------------------	-------------------

Jurnal Pendidikan Ekonomi dan Akuntansi

Vol. 3, No. 2, Juni 2010

ISSN 1978-8770

Jurnal Pendidikan Ekonomi dan Akuntansi (JPEA) merupakan wadah informasi bidang pendidikan ekonomi dan akuntansi berupa publikasi hasil karya penelitian, karya ilmiah lain yang relevan, dan studi kepustakaan.

Jurnal ini terbit pertama kali pada Bulan Desember 2007 dengan frekuensi penerbitan dua kali dalam setahun pada Bulan Juni dan Desember.

Dewan Redaksi

Pemimpin Umum : L. Saptono, S.Pd., M.Si.

Sekretaris Redaksi : B. Indah Nugraheni, S.I.P., M.Pd.

Anggota : S. Widanarto Prijowuntato, S.Pd., M.Si.
E. Catur Rismiati, S.Pd., M.A.
Drs. FX. Muhadi, M.Pd.

Redaktur Ahli : Dr. Fr. Ninik Yudianti, M.Acc.
(Universitas Sanata Dharma Yogyakarta)
Prof. Dr. Kisyani Laksono, M.Hum.
(Universitas Negeri Surabaya)
Prof. Dr. Augusty Ferdinand, M.B.A.
(Universitas Diponegoro Semarang)
Dr. C. Asri Budiningsih
(Universitas Negeri Yogyakarta)

Tata letak : Ig. Bondan Suratno, S.Pd., M.Si.

Administrasi : Natalina Premastuti B., S.Pd., M.Pd.

Alamat Redaksi

Program Studi Pendidikan Akuntansi, FKIP Universitas Sanata Dharma

Jl. Affandi, Mrican, Tromol Pos 29 Yogyakarta 55002. Telp (0274) 513301 Ext. 1445 Faks. (0274) 562383

E-mail: prodipakusd@staff.usd.ac.id

Berlangganan

Hubungi Bagian Tata Usaha Lembaga Penelitian dan Pengabdian Kepada Masyarakat USD

Jl. Affandi, Mrican, Tromol Pos 29 Telp. (0274) 513301 Ext. 1527, Faks. (0274) 540793

E-mail: lemlit@staff.usd.ac.id.

Harga per exemplar Rp 20.000,-.

KARAKTERISTIK PERANGKAT TES TEORI KEJURUAN SMK BISNIS DAN MANAJEMEN PROGRAM KEAHLIAN AKUNTANSI DI DAERAH ISTIMEWA YOGYAKARTA

Natalina Premastuti Brataningrum^{*)}

Abstract

The objectives of the research were: 1) to describe the qualitative characteristics of test instrument on accountancy theory of vocational competence examination on 2008/2009 study year in DIY province; 2) to describe the qualitative characteristics of the test instrument on accountancy theory of vocational competence examination on 2008/2009 study year in DIY province.

This was an avaluative descriptive research. The respondent of this research were 2799 students that follow the Nasional Examination on Accounting theory test subject. Besides the students' respond data, this research also needs the test questions document of vocational theory. The data analysis is done qualitatively by consulting to the expert in précising the question according to the arranged criteria. On the other hand, the quantitative analysis is done using the classical theory approach and Rasch Model.

The result of the this research were; 1) the characteristic of 40 question items on Accountancy theory that are based on qualitative analysis by item analyzing show 57,5% are on good category and 42,5% are bad, 2) based on qualitative analysis using classical test theory, there are 80% good items and 10% bad items with 0,779 realibility index 0,779 so they belong to good category. The good question items based on item respond theory are 77,5% and bad question item are 22,5% with maximum test information function 14,5 on $\theta = -1,0$

Kata Kunci: *characteristic of test instrument, vocational competence examination, National Examination*

A. Pendahuluan

Latar Belakang Masalah

Undang-Undang RI Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional BAB XVI pasal 57 ayat (1) menyatakan bahwa evaluasi dilakukan dalam rangka pengendalian mutu pendidikan secara nasional sebagai bentuk akuntabilitas penyelenggaraan pendidikan kepada pihak-pihak yang berkepentingan. Evaluasi dilakukan untuk menjaga agar kualitas penyelenggaraan pendidikan sesuai dengan kebutuhan masyarakat secara terus menerus. Dengan demikian, hasil evaluasi pendidikan dapat digunakan sebagai dasar untuk menyusun kebijakan sesuai dengan tuntutan pendidikan yang berubah sangat cepat.

Penyelenggaraan evaluasi adalah untuk mendapatkan gambaran yang konkrit

tentang kompetensi peserta didik. Salah satu variabel yang perlu diperhatikan adalah bentuk alat ukur yang dipakai. Teknik-teknik penyusunan tes yang digunakan diharapkan memberikan landasan yang kuat untuk melakukan evaluasi yang tepat. Hasil pengukuran yang diperoleh harus mempunyai tingkat kesalahan sekecil mungkin. Tingkat kesalahan erat kaitannya dengan alat ukur yang dipakai. Terdapat dua kesalahan yang ditemui dalam alat ukur, yakni kesalahan yang bersifat acak dan kesalahan yang bersifat sistematis (Djemari Mardapi, 2007: 3). Kesalahan acak disebabkan kondisi fisik dan mental pada pihak yang diukur maupun pihak yang mengukur. Sebagai contohnya, peserta tes yang memiliki tingkat kecemasan yang tinggi apabila di tes maka besar kemungkinan hasil pengukuran tidak menunjukkan hasil kemampuan sesungguhnya. Berbeda dengan

^{*)} Natalina Premastuti Brataningrum adalah staf pengajar pada Program Studi Pendidikan Akuntansi, USD

kesalahan acak, kesalahan sistematis disebabkan oleh alat ukurnya, yang diukur, dan yang mengukur. Sebagai contohnya, pendidik memberikan soal terlalu mudah atau terlalu sulit. Tindakan pendidik ini akan berimplikasi pada hasil pengukuran yang bisa di bawah atau di atas kemampuan peserta didik yang sebenarnya.

Berbagai fenomena terkait dengan kesalahan instrumen tes ujian kian merebak. Sebagai contohnya adalah kesalahan dalam penulisan soal ataupun kunci jawaban yang lebih dari satu jawaban benar. Bahkan dalam beberapa kasus terdapat soal yang tidak termasuk standar kompetensi yang disyaratkan pada level tersebut. Kesalahan-kesalahan tersebut sangat mungkin dilakukan oleh siapapun sekalipun instrumen telah disusun melalui mekanisme yang tepat.

Sekolah Menengah Kejuruan (SMK) merupakan institusi pendidikan formal yang bertujuan menyelenggarakan pendidikan bagi peserta didik agar memiliki kemampuan akademik maupun *skills* yang di kemudian hari siap untuk memasuki dunia kerja. Sejalan dengan tujuan tersebut, pembelajaran diselenggarakan di sekolah maupun di dunia usaha. Sehubungan dengan tujuan tersebut, maka bentuk pengujian yang dilakukan tidak hanya meliputi aspek teori, tetapi juga pada aspek praktik. Sistem pengujian pada aspek praktik dan teori ini telah lama diterapkan pada peserta didik SMK. Pada tahun ajaran 2008/2009, ujian teori yang diselenggarakan program studi akuntansi menggunakan bentuk tes, yakni penyajian soal dengan memberikan serentetan pertanyaan dengan tipe pilihan ganda dengan jumlah soal sebanyak 40 butir. Penelitian ini dimaksudkan untuk mengungkapkan karakteristik instrumen tes ujian teori akuntansi pada SMK Bisnis dan Manajemen. Karakteristik instrumen tes ujian yang dimaksud meliputi karakteristik kuantitatif dan kualitatif.

Batasan Masalah

Penelitian ini hanya terbatas pada pengungkapan karakteristik perangkat tes Ujian Kompetensi Kejuruan Teori Akuntansi SMK Bisnis dan manajemen program studi

akuntansi tahun ajaran 2008/2009 di wilayah Propinsi DIY.

Rumusan Masalah

Rumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana karakteristik kualitatif perangkat tes Ujian Kompetensi Kejuruan Teori Akuntansi tahun ajaran 2008/2009 di wilayah Propinsi DIY?
2. Bagaimana karakteristik kuantitatif perangkat tes Ujian Kompetensi Kejuruan Teori Akuntansi tahun ajaran 2008/2009 di wilayah Propinsi DIY?

Tujuan Penelitian

Berdasarkan rumusan masalah tersebut di atas, maka tujuan penelitian ini adalah sebagai berikut:

1. Mendeskripsikan karakteristik kualitatif perangkat tes Ujian Kompetensi Kejuruan teori akuntansi tahun ajaran 2008/2009 di Propinsi DIY.
2. Mendeskripsikan karakteristik kuantitatif perangkat tes Ujian Kompetensi Kejuruan teori akuntansi tahun ajaran 2008/2009 di Propinsi DIY.

Manfaat Penelitian

1. Sebagai masukan bagi guru/pengembang tes terkait tentang prosedur penyusunan dan kriteria tes yang baik, serta alternatif metode analisis butir tes yang paling tepat.
2. Bagi sekolah dapat digunakan sebagai informasi untuk tindak lanjut pengambilan keputusan dalam penyusunan dan pengembangan tes, serta evaluasi pembelajaran di sekolah.

B. Kajian Teori

Karakteristik Mata Pelajaran Akuntansi

Mata pelajaran Akuntansi merupakan bagian dari materi produktif yang harus dikuasai oleh peserta didik SMK Bisnis dan Manajemen program keahlian akuntansi. Mata pelajaran akuntansi di dalam kurikulum telah disajikan mulai dari kelas 1 sampai kelas 3. Muatan kompetensi yang harus

dikuasai pada mata pelajaran ini telah sejalan dengan tujuan pendidikan SMK, yakni menyiapkan peserta didik sehingga memiliki kemampuan sebagai tenaga keuangan tingkat menengah.

Horngren dan Harrison (1989: 2) mendefinisikan akuntansi sebagai sebuah sistem yang digunakan untuk mengukur aktivitas bisnis, memproses informasi sehingga dapat disajikan menjadi laporan, dan mengkomunikasikan berbagai temuan tersebut bagi penentu kebijakan. Berdasarkan definisi tersebut dapat disimpulkan bahwa mata pelajaran akuntansi membahas berbagai hal terkait dengan proses identifikasi bukti transaksi, pencatatan transaksi, *posting* ke dalam buku besar, neraca saldo, jurnal penyesuaian, pembuatan neraca lajur, dan penyusunan laporan keuangan. Pada intinya materi yang disajikan ditujukan dalam rangka untuk mengetahui kinerja perusahaan, baik perusahaan jasa dan perusahaan dagang.

Evaluasi Hasil Belajar

Kegiatan evaluasi belajar merupakan proses yang dilakukan untuk mengetahui ketercapaian kompetensi peserta didik. Selanjutnya berdasarkan informasi tersebut dapat dijadikan sebagai umpan balik untuk mengetahui ketercapaian tujuan pembelajaran yang telah ditetapkan. Menurut Gronlund dan Linn (1990: 5), *evaluation is systematic process of collecting, analyzing, and interpreting information to determine the extent to which pupils are achieving instructional objectives*. Sejalan dengan pendapat Gronlund dan Linn tersebut, Klizlik (2009: 2) mengatakan bahwa evaluasi adalah proses yang didesain untuk mendapatkan informasi dan dapat digunakan sebagai bahan pertimbangan untuk membuat keputusan (<http://www.adprima.com/measurement.htm>).

Evaluasi hasil belajar tidak terlepas dari kegiatan pengukuran dan penilaian, artinya kedua kegiatan tersebut tidak dapat dipisahkan satu sama lain, serta dilaksanakan secara berurutan. Griffin dan Nix (1991: 3) menyatakan bahwa pengukuran, penilaian, dan evaluasi adalah sebuah hirarki. Pengukuran membandingkan hasil

pengamatan dengan kriteria, penilaian menjelaskan dan menafsirkan hasil pengukuran, sedangkan evaluasi adalah penetapan nilai dan implikasi perilaku. Pengukuran adalah membandingkan sesuatu dengan alat ukurnya dan merupakan kegiatan penentuan bilangan bagi obyek secara sistematis. Menurut Ebel dan Frisbie (1986: 14), *measurement is the process of assigning numbers to individuals or their characteristics according to specified rules*.

Penilaian merupakan langkah lanjutan setelah dilakukan pengukuran. Informasi yang diperoleh dari hasil pengukuran selanjutnya dideskripsikan dan ditafsirkan. Penafsiran terhadap hasil pengukuran ini didasarkan pada suatu kriteria. Kriteria yang digunakan dalam penilaian hasil belajar dapat dilakukan dengan menggunakan acuan norma dan acuan kriteria (Sax, 1980: 22).

Tes Prestasi Belajar

Instrumen pengukuran yang digunakan dalam penilaian hasil belajar dapat berupa tes prestasi belajar. Syaifudin Azwar (1999: 9) mengatakan, tes prestasi merupakan tes yang disusun secara terencana untuk mengungkap kemampuan maksimal subyek, dalam menguasai bahan-bahan atau materi yang diajarkan. Hal senada dikemukakan Allen dan Yen (1979: 1) yang menyebutkan bahwa *a test is device for obtaining a sample of an individual's behavior*. Artinya, tes sebagai instrumen atau prosedur sistematis untuk mengukur perilaku individu.

Ebel dan Frisbie (1986: 20-21) mengungkapkan beberapa fungsi tes prestasi yakni: (1) mengukur prestasi belajar siswa; (2) mempunyai kontribusi terhadap program pengajaran; dan (3) sebagai motivasi terhadap belajar siswa. Agar diperoleh informasi data yang akurat, tes yang digunakan harus memiliki bukti kesahihan dan handal, hal ini bertujuan untuk meningkatkan kualitas pendidikan (Mardapi, 2008: 7). Persyaratan tes yang baik dalam konsep pengukuran, menurut Mehrens dan Lehman (1973: 101) adalah sebagai berikut *perhaps the two most important technical concepts in measurement are validity and reliability*. Senada dengan hal tersebut, Li (2003: 90) mengatakan bahwa reliabilitas dan

validitas merupakan prinsip dasar dalam psikometri, ini berarti bahwa tes sebagai alat ukur hasil belajar harus memenuhi syarat prinsip dasar pengukuran yaitu tes yang valid dan reliabel.

Bentuk Tes

Tes yang diselenggarakan dapat menggunakan berbagai bentuk. Gronlund (1976: 144) menjelaskan bahwa berdasarkan bentuk pertanyaannya, tes ada dua macam yaitu tes bentuk obyektif dan esai. Tes obyektif adalah tes yang keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia. Sejalan dengan pendapat tersebut Djemari Mardapi (2008: 70) mengatakan, bentuk tes obyektif yang sering digunakan adalah bentuk pilihan ganda, benar salah, menjodohkan, dan uraian obyektif. Tes pilihan ganda merupakan perangkat tes yang terdiri dari pernyataan atau pertanyaan (*stem*), beberapa pilihan jawaban, kunci jawaban, dan pengecoh jawaban atau pengecoh. Menurut Gronlund dan Linn (1990: 166-167):

multiple choice item consists of problem and list of suggested solutions. The problem may be stated as direct question or an incomplete statement and is called the stem of item. The list of suggested solutions may include words, numbers, symbol, or phrases are called alternatives (also called options)...The correct alternatives in each item is called merely the answer, and the remaining alternatives are called distracters.

Menurut Syaifudin Azwar (1999: 57), dari segi materi tes harus komprehensif dan berisi butir-butir yang relevan. Komprehensif artinya mencakup keseluruhan isi atau bahan pelajaran yang telah diidentifikasi sebagai tujuan pengukuran, relevan artinya butir-butir tes yang akan dituliskan berkaitan dan dianggap perlu untuk memahami materi.

Pada umumnya penulisan item pilihan ganda didasarkan pada proposisi yang dikembangkan sedemikian rupa sehingga jawabannya tidak sekedar menjawab benar salah, tetapi berupa pilihan terhadap pernyataan yang paling benar (Azwar, 1999: 81). Untuk itu terdapat beberapa pedoman

yang dapat digunakan untuk penulisan tes pilihan ganda yang efektif.

Dalam menulis soal pilihan ganda juga sangat diperlukan kemampuan menerjemahkan gagasan-gagasan dalam bahasa yang komunikatif. Soal harus ditulis dalam bahasa yang sederhana dan jelas, yaitu bahasa yang terbebas dari isi yang tidak fungsional dan bebas dari sumber-sumber kesulitan yang tidak relevan (misalnya, ambiguitas) yang dapat menyebabkan siswa yang tahu justru salah menjawab. Menurut Popham (1995: 99), salah satu ketentuan penulisan tes yang baik adalah menghindari menuliskan pertanyaan yang bersifat ambigu, oleh sebab hal ini dapat menimbulkan salah persepsi bagi siswa, sehingga akan memberikan jawaban yang salah. Soal juga harus bebas dari petunjuk yang tidak relevan (misalnya, asosiasi verbal), oleh sebab dapat menyebabkan siswa yang tidak tahu justru menjawab dengan benar. Selain itu, format soal harus memungkinkan tanggapan yang efisien dan mengikuti aturan tata bahasa yang baku. Menurut Sumadi Suryabrata (2005: 75), penulis soal harus sadar akan kemungkinan adanya arti ganda kata-kata atau makna rangkap dalam kalimat yang digunakan dalam soal. Gagasan yang merupakan isi soal harus diterjemahkan dalam bahasa yang jelas dan mudah dimengerti.

Hal yang juga perlu diperhatikan dalam penulisan soal pilihan ganda adalah menentukan batasan perilaku atau kompetensi, yaitu kemampuan yang akan diukur sesuai dengan tujuan instruksional yang telah ditetapkan. Untuk merencanakan tingkat penguasaan kompetensi dapat dilakukan berdasarkan sasaran kompetensi kognitif yang diharapkan dalam tes, yang mengacu pada taksonomi Bloom. Taksonomi disusun dari level kognisi yang paling sederhana, yaitu ingatan (C1) hingga yang paling kompleks yaitu evaluasi (C6). Gronlund (1990: 32) mengatakan bahwa *the major categories in the cognitive domain are knowledge, comprehension, application, analysis, synthesis and evaluation. These categories begin with the relatively simple knowledge outcomes and proceed through increasingly complex levels of intellectual ability.* Menurut Syaifudin Azwar (1999: 63),

taksonomi kognitif Bloom disusun sedemikian rupa sehingga kompleksitasnya bertambah secara bertingkat dari taraf *knowledge* yang paling rendah ke taraf *evaluation* yang paling tinggi. Konsekuensinya, semakin tinggi kemampuan yang diukur sesuai dengan target kompetensi, maka semakin sulit soal dan semakin sulit pula menyusunnya. Keseluruhan butir dalam tes yang direncanakan terdiri atas beberapa taraf kompetensi yang berberda-beda, sudah barang tentu untuk menentukan taraf kesukaran berdasarkan taksonomi ini harus memperhatikan urgensi pencapaian kompetensi berdasarkan materi.

Meskipun memiliki banyak kelebihan, penggunaan tes pilihan ganda tetap mendapatkan kritik karena memiliki kelemahan-kelemahan. Menurut Nitko dan Bookhart (1996: 152), jika tes pilihan ganda digunakan eksklusif untuk penilaian *highstakes* mungkin akan menghasilkan model pendidikan yang tidak dikehendaki. Misalnya penilaian pada ujian yang menggunakan tes pilihan ganda yang fokus pada pengetahuan faktual, para guru akan cenderung menempuh teknik *drill* dan praktis dalam mempersiapkan siswanya.

Teori Tes Klasik

Penggunaan teori tes klasik dalam kontruksi dan intepretasi skor telah berjalan beberapa dekade lamanya. Djemari Mardapi (2008: 32) mengungkapkan bahwa teori tes klasik menggunakan model pengukuran yang sangat sederhana, yakni skor yang tampak merupakan penjumlahan dari skor murni dan kesalahan pengukuran. Hubungan antara skor tampak (X), skor murni (T) dan kesalahan pengukuran (E) dapat ditulis dalam persamaan: $X = T + E$.

Ada tujuh macam asumsi teori klasik. Menurut Allen & Yen (1979: 57-60) ketujuh asumsi teori klasik tersebut yaitu: (1) terdapat hubungan antara skor tampak (*observed score*) yang dilambangkan dengan huruf X, skor murni (*true score*) yang dilambangkan dengan T dan skor kesalahan (*error*) yang dilambangkan dengan E; (2) skor murni (T) merupakan nilai harapan X (*expected value of X*) yaitu $E(X)$, jadi T merupakan harga rerata distribusi teoritik skor X apabila orang yang sama dikenai tes

yang sama berulang kali dengan asumsi pengulangan tes itu dilakukan tidak terbatas banyaknya dan setiap pengulangan tes adalah independen satu sama lain; (3) tidak ada korelasi antara *error* pengukuran dan skor murni; (4) korelasi antar kesalahan pada pengukuran pertama dan pada pengukuran kedua adalah ($=0$); (5) jika terdapat dua tes untuk mengukur atribut yang sama maka skor kesalahan pada tes pertama (E_1) tidak berkorelasi dengan skor murni pada tes kedua (T_2); (6) menyajikan tentang pengertian tes yang paralel; (7) definisi tes yang setara (*essentially equivalent*). Jika dua perangkat tes mempunyai skor-skor perolehan dan yang memenuhi asumsi 1 sampai 5 dan apabila untuk setiap populasi subyek $X_1 = X_2 + C_{12}$, dimana C_{12} adalah sebuah bilangan konstanta, maka kedua tes itu disebut tes yang paralel.

Asumsi-asumsi teori klasik sebagaimana disebutkan sebagai dasar dikembangkannya berbagai formula yang berguna untuk mengetahui karakteristik perangkat tes yaitu reliabilitas, validitas, tingkat kesukaran, daya beda, dan kesalahan baku pengukuran.

Model Rasch

Model Rasch pertama kali diajukan oleh George Rasch sekitar tahun 1966 yang kemudian dikembangkan di Amerika oleh Benjamin Wright di Universitas Chicago. Parameter butir pada model ini adalah tingkat kesukaran butir, sedangkan parameter lainnya seperti daya pembeda dianggap sama, dan dugaan atau *guessing* sama dengan nol. Berikut ini akan disajikan formula model Rasch (Wright & Stone, 1979: 16).

$$\pi_{vi} = \frac{\exp(\beta_v - \delta_i)}{[1 + \exp(\beta_v - \delta_i)]}$$

π adalah peluang peserta tes yang memiliki kemampuan v menjawab butir i dengan benar, β_v adalah tingkat kemampuan peserta tes, parameter δ adalah tingkat kesulitan butir i , dan \exp adalah bilangan transendental yang memiliki nilai 2, 718. Ketika peserta tes v lebih memiliki kemampuan daripada tingkat kesulitan item i , kemudian β_v lebih besar dari δ_i , maka

akan terjadi selisih positif dan probabilitas untuk menjawab item i lebih dari 0,5. Jika kemampuan seseorang jauh melebihi tingkat kesulitan item, maka akan terjadi selisih positif yang lebih besar, yang pada akhirnya probabilitas menjawab benar mendekati 1. Tetapi ketika item terlalu sukar dijawab oleh peserta tes, kemudian $\hat{\alpha}_i$ lebih kecil dari $\hat{\alpha}_j$, selisihnya menghasilkan angka negatif sehingga probabilitas peserta tes untuk menjawab benar kurang dari 0,5. Selanjutnya, ketika tingkat kesukaran item jauh melampaui kemampuan seseorang, maka probabilitas peserta tes menjawab benar akan mendekati nol.

Tingkat kesukaran pada model *Rasch* merupakan fungsi dari kemampuan seseorang. Seseorang yang mempunyai kemampuan tinggi akan merasa mudah mengerjakan butir soal, sebaliknya seseorang yang memiliki kemampuan rendah akan sulit menjawab butir soal. Asumsi yang digunakan pada model *Rasch* adalah: (1) semua butir memiliki daya pembeda yang sama, dan (2) peluang menjawab butir benar bagi yang memiliki kemampuan rendah sama dengan nol. Dengan kata lain semua kurva karakteristik butir soal adalah sejajar atau mendekati sejajar, karena besarnya slope kurva sama dan memotong sumbu y di titik 0 atau mendekati 0.

C. Metode Penelitian

Jenis Penelitian

Penelitian ini merupakan penelitian yang menggunakan pendekatan deskriptif evaluatif. Dalam penelitian ini akan diungkapkan karakteristik butir tes ujian teori akuntansi baik dari indeks kesukaran, indeks pembeda, pengecoh, maupun reliabilitasnya yang kemudian akan dinilai apakah telah sesuai dengan kriteria ideal.

Pengambilan Data dan Waktu Penelitian

Pengambilan data respon peserta tes teori akuntansi dilaksanakan di Kantor Pusat Penilaian Pendidikan Balitbang Depdiknas Jakarta. Teknik pengambilan data ini menggunakan teknik dokumentasi. Selanjutnya instrumen yang digunakan untuk melakukan telaah kualitatif dikonsultasikan kepada ahli/pakar.

Populasi Penelitian

Jumlah populasi dalam penelitian ini adalah sebesar 2.799 peserta didik (peserta tes).

Subjek dan Objek Penelitian

Subjek penelitian ini adalah mata pelajaran akuntansi jenjang SMK tahun

Tabel: Hasil Penelaahan Butir Soal Akuntansi Berdasarkan Analisis Kualitatif

Aspek	Kriteria	Butir yang tidak memenuhi kriteria	
		Penelaah 1	Penelaah 2
Materi	1		
	2	5,11,23,29	8,1
	3	14,17	17
	4		
	5	17	17
Konstruksi	1		
	2		8,10,11
	3		
	4		
	5		
	6		33
	7	29	8
	8		
Bahasa	1	27,30,37	
	2	9,17,35	17
	3		
	4		
	5		

pelajaran 2008/2009. Objek penelitian ini adalah respon peserta ujian, dan lembar soal ujian nasional teori kejuruan program keahlian akuntansi jenjang SMK tahun ajaran 2008/2009.

Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan dengan pendekatan kualitatif dan kuantitatif. Pendekatan kualitatif berupa analisis data melalui penelaahan butir soal. Telaah butir soal dilakukan dari aspek materi, aspek konstruksi, dan aspek bahasa, berdasarkan kriteria penelaahan yang ditetapkan oleh Pussajian Balitbang Depdiknas dalam bentuk format kartu telaah sebagai instrumen utama. Pendekatan kuantitatif dilakukan melalui analisis empiris terhadap butir soal yang direspon oleh peserta tes. Pada penelitian ini digunakan analisis butir soal menurut teori tes klasik, dengan program *Microcat Ite-man* versi 3.00 dan analisis butir soal menurut model *Rasch* dengan program *Microcat Bigstep* versi 2.30, dengan tujuan untuk saling melengkapi.

D. Hasil Penelitian dan Pembahasan

Hasil Penelitian

Hasil Analisis Kualitatif

Berdasarkan analisis secara kualitatif seperti tampak pada tabel 1 di atas, dapat disimpulkan bahwa butir soal teori akuntansi termasuk dalam kategori cukup baik. Hal ini dikarenakan 57,5% butir masuk dalam kriteria baik sementara 42,5% termasuk dalam kriteria tidak baik.

Hasil Analisis Kuantitatif dengan Pendekatan Teori Tes Klasik

Tingkat Kesukaran

Berdasarkan hasil analisis dengan program *Ite-man*, tingkat kesukaran butir soal Akuntansi dapat ditunjukkan pada Tabel 3 berikut ini.

Dari tabel 2 di bawah dapat dilihat bahwa terdapat 3 (7,5%) butir soal masuk dalam kategori sukar, 12 (30%) butir soal masuk dalam kategori sedang, dan 25 (62,5%) butir soal masuk dalam kategori mudah.

Tabel 2: Tingkat Kesukaran Butir Soal Akuntansi

Kategori	Nomor Butir Soal	Jumlah	Persentase
$p > 0,75$ Mudah	1,2,3,5,6,7,10,11,15,16,18,19,21,23,25,26,27,29,30,32,35,36,37,39,40.	25	62,5
$0,25 \leq p \leq 0,75$ Sedang	4,12,13,14,20,22,24,28,31,33,34,38	12	30
$p < 0,25$ Sukar	8,9,17,	3	7,5
Jumlah		40	100

Kesimpulan yang dapat ditarik adalah sebagian besar butir soal tes tersebut dikategorikan mempunyai tingkat kesukaran yang sedang.

Daya Beda

Hasil analisis daya beda butir soal teori akuntansi dapat dilihat pada Tabel 4 sebagai berikut.

Dari tabel 3 di bawah, dapat dilihat bahwa 31 (77,5%) butir soal mempunyai daya beda yang baik dan 9 (22,5%) butir soal mempunyai daya beda yang tidak baik.

Efektifitas Pengecoh

Hasil analisis efektifitas pengecoh butir soal akuntansi menggunakan program *Ite-man* disajikan pada Tabel 4 di bawah ini.

Dari hasil analisis butir soal matematika yang disajikan pada tabel 4 di bawah, nampak 33 (82,5%) butir soal memiliki pengecoh yang baik, karena memiliki r_{pbis} negatif. Sedangkan pengecoh yang tidak baik sebanyak 7 (17,5%) butir soal, karena memiliki positif selain kunci. Selain itu terdapat 3 butir soal dengan peringatan "check the key", yaitu nomor 8,9 dan 17.

Tabel 3: Daya Beda Butir Soal Akuntansi

Kategori	Nomor Butir Soal	Jumlah	Persentase
>0,20 Baik	1,2,3,4,5,6,7,10,11,12,13,14,15,16,18,19, 20,21,22,23,24,25,26,27,28,29,30,31,32, 33,34,35,36,37,38,39,40	37	92,5
<0,20 Tidak baik	8,9,17	3	7,5
Jumlah		40	100

Tabel 4: Efektifitas Pengecoh Butir Soal Akuntansi

Kategori	Nomor Butir Soal	Jumlah	Persentase
r_{pbis} negatif (Baik)	1,2,3,4,5,6,7,10,11,12,13,14,15,16 18,19,20,21,23,24,25,26,27,28,29 30,31,32,33,35,37,39,40	33	82,5
r_{pbis} positif selain kunci jawaban (Tidak baik)	8,9,17,22,34,36,38	7	17,5
Jumlah		40	100

Skala Statistik

Tabel 5: Skala Statistik Tes Akuntansi

No	Skala Statistik	Hasil	No	Skala Statistik	Hasil
1.	Jumlah butir soal	40	9.	Skor maksimum	37
2.	Jumlah peserta tes	2799	10.	Median	28
3.	Rerata	27,503	11.	Alpha	0,779
4.	Varian	23,218	12.	SEM	2,267
5.	Simpangan baku	4,818	13.	Mean P	0,688
6.	Skew	-0,718	14.	Mean item-tot	0,314
7.	Kurtosis	0,327	15.	Mean biserial	0,471
8.	Skor minimum	6			

Hasil Analisis Kuantitatif dengan Model Rasch

Tabel 6: Statistik Tes Bahasa Indonesia dengan Program Bigsteps

Besaran Statistik	Peserta Tes	Butir Soal
Rerata skor	22,8	1968,5
Standar deviasi skor	3,9	758,4
Rerata kemampuan	1,39	0,00
Standar deviasi kemampuan	1,45	2,86
Rerata kesalahan pengukuran	1,03	0,09
Standar deviasi kesalahan pengukuran	1,34	0,1

Kecocokan dengan Model

Hasil analisis kecocokan model butir soal ujian nasional teori akuntansi disajikan pada tabel 7 di bawah ini.

Berdasarkan hasil analisis dengan pro-

gram *Bigsteps* pada Tabel 7 tersebut, diperoleh informasi bahwa butir soal yang cocok dengan model (*fit*) sebanyak 31 (77,5%). Sedangkan butir soal yang tidak cocok dengan model sebanyak 9 (22,5%)

Tabel 7: Kecocokan Butir Soal dengan Model Rasch

Kriteria	Nomor Butir	Jumlah Butir	Persentase
Oufit < 2,00 dan Ptbis positif (cocok dengan model)	1,2,3,6,7,8,9,10,11,12,13,15,16,17,18,19,21,23,25,26,27,28,29,30,32,33,35,37,38,39,40.	31	77,5
Oufit > 2,00 dan Ptbis negatif (tidak cocok dengan model)	4,5,14,20,22,24,31,34,36.	9	22,5
Jumlah		40	100

Tabel 8: Distribusi Tingkat Kesukaran Butir Soal Ujian Nasional Teori Akuntansi yang Cocok dengan Model Rach

Kriteria	Nomor Butir	Jumlah	Persentase
Sukar (>2,00)	8,9,13,17	4	12,9
Sedang (-2,00 s/d 2,00)	1,2,3,6,7,10,11,12,15,16,18,19,21,23,25,26,27,28,29,30,33,35,37,38,39,40	26	83,87
Mudah (< -2,00)	32	1	3,23
Jumlah		31	100

Berdasarkan tabel 8 di atas diperoleh informasi bahwa dari 31 butir soal yang cocok dengan model, terdapat butir soal dengan tingkat kesukaran yang baik sebanyak 26 (83,87%). Butir soal nomor 8,9,13,17 tergolong butir soal sukar dan butir nomor 32 mudah sehingga tergolong tidak baik.

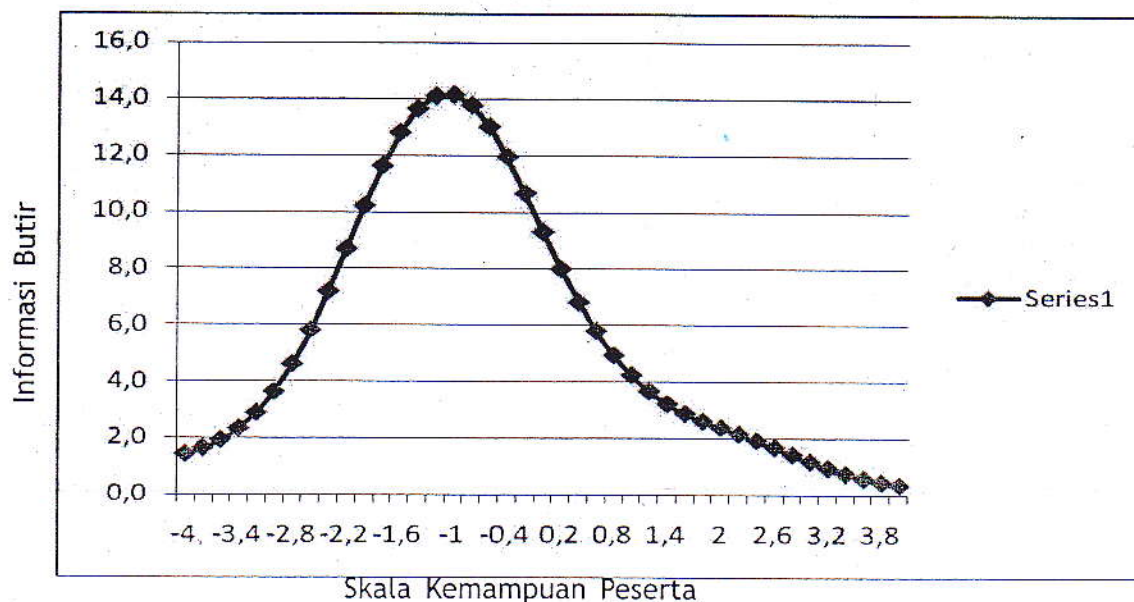
Fungsi Informasi

Berdasarkan penghitungan nilai informasi butir, semua butir memiliki nilai informasi

lebih dari 0,7. Selanjutnya nilai fungsi informasi tes maksimal sebesar 14,15 dan nilai ini tercapai jika parameter kemampuan peserta sebesar -0,1. Grafik nilai informasi tes beserta harga *standard error of estimation* pada rentang -4,0 4,0.

Selanjutnya dapat dilihat lengkungan fungsi informasi tes Teori Akuntansi SMK Bisnis dan Manajemen di DIY tahun ajaran 2008/2009 seperti pada Gambar 5 di bawah.

Gambar 2: Kurve Fungsi Informasi Tes Ujian Nasional Teori Akuntansi



B. Pembahasan

Hasil Analisis Kualitatif

Secara teoritis, hasil telaah kualitatif terhadap butir-butir yang terdapat pada perangkat tes teori akuntansi SMK Bisnis dan Manajemen di DIY tahun ajaran 2008/2009, sebagian besar butir soal sudah masuk dalam kategori baik. Hal ini terlihat dari terpenuhinya hampir semua kriteria pada aspek materi, konstruksi dan bahasa. Namun demikian masih terdapat beberapa soal yang belum memenuhi kriteria.

Pada aspek materi terdapat 13 butir yang tidak memenuhi kriteria. Kriteria pengecoh tidak dipenuhi oleh 6 butir yakni butir 1,5, 8,11, 23, 29. Sementara kriteria kunci jawaban tidak dipenuhi oleh butir 17. Soal No. 17 terkait dengan topik buku besar, pilihan yang dianggap sebagai jawaban adalah E dimana bunyi pilihan jawabannya adalah media untuk memposting transaksi dari beragam akun ke dalam satu akun. Pilihan ini menjadi tidak tepat karena belum memberikan informasi yang memadai. Untuk menjawab pertanyaan no.17, pilihan jawaban C memberikan informasi yang sangat memadai dan lebih tepat berfungsi sebagai kunci jawaban. Pada kriteria pilihan jawaban benar-benar berfungsi terdapat satu nomor yakni no.17 yang tidak memenuhi kriteria.

Pada aspek konstruksi terdapat 11 item yang tidak memenuhi kriteria yang disyaratkan. Pokok (stem) soal dan jawaban pada butir 8,10,11 tidak dirumuskan dengan tegas. Butir 33 tidak memenuhi kriteria pilihan jawaban, angka seharusnya diurutkan berdasarkan besar kecilnya angka. Untuk kriteria pilihan jawaban homogen, butir 8 dan 29 tidak memenuhi persyaratan tersebut. Pada aspek bahasa terdapat 6 butir yang tidak memenuhi kriteria. Butir 27, 30, 37 tidak memenuhi kriteria bahwa kalimat tidak ditulis dengan menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya. Butir 9,17,35 tidak memenuhi kriteria bahwa kalimat tidak menimbulkan penafsiran ganda.

Hasil Analisis Kuantitatif

Hasil Analisis Teori Tes Klasik dengan Program *Iteman*

Berdasarkan hasil analisis teori tes klasik dengan program *Iteman* dapat dikemukakan beberapa pembahasan sebagai berikut.

Dari hasil analisis tingkat kesukaran, dapat dilihat bahwa sebanyak 12 (30%) butir soal masuk dalam kategori sedang, artinya butir soal tersebut sudah berfungsi dengan baik. Sedangkan 28 (70%) butir soal belum berfungsi dengan baik, karena butir soal tersebut masuk kategori terlalu mudah dan terlalu sukar.

Dilihat dari hasil analisis daya beda, dapat dikatakan bahwa hampir seluruh butir Akuntansi berfungsi dengan cukup baik, karena sebanyak 37 butir soal atau 92,5% memenuhi persyaratan yang ditentukan, artinya sebagian besar butir soal telah dapat membedakan kemampuan antara siswa yang mempunyai kemampuan tinggi dan rendah.

Ditinjau dari distribusi jawaban, maka dapat dijelaskan bahwa sebagian besar butir masuk kategori baik. Butir soal yang baik berdasarkan distribusi jawaban tes Akuntansi sebanyak 33 butir soal atau 82,5%. Hal ini menunjukkan bahwa pengecoh telah berfungsi dengan baik, karena sebagian besar pengecoh sudah dipilih oleh siswa yang mempunyai kemampuan rendah.

Hasil analisis butir soal Akuntansi dengan menggunakan *Iteman* diperoleh skala statistik sebagai berikut. Rerata skor tampak peserta adalah 27,503 dengan nilai simpangan baku sebesar 4,818 dan rentang skor dari 6 sampai dengan 37. Harga skew sebesar -0,718, tanda negatif bermakna distribusi menceng ke kiri hal ini dapat dilihat dari harga rerata yang lebih kecil dari median ($27,503 < 28$) Rerata tingkat kesukaran sebesar 0,688 menunjukkan rerata jumlah butir soal yang dijawab benar oleh siswa menunjukkan tingkat kesukaran yang sedang. Dari hasil analisis juga diperoleh koefisien reliabilitas sebesar 0,779 termasuk perangkat tes yang sudah reliabel. Nilai indeks reliabilitas berkaitan dengan *standar error of measurement*. Dari hasil analisis diperoleh indeks sebesar 2.267, besaran ini menunjukkan tingkat ketelitian hasil pengukuran.

- Leydens, J. A., Moskal, B. M. (2000). Scoring Rubric Development: Validity and Reliability. *Practical Assessment, Research & Evaluation*. Diambil 15 Nopember 2009. dari <http://pareonline.net/getvn.asp?v=7&n=10>.
- Li, H. (2003). The Resolution of Some Paradoxs Related to Reliability and Validity. *Journal of Educational and Behavioral Statistics*, Vol.28, No.2, pp. 89-95.
- Linn, R. L. (1989). *Education Measurement*. New York: Macmilan Publishing Company
- Mardapi, Djemari. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendekia.
- Mehrens, W. A., Lehmann, I. J. (1973). *Measurement and Evaluation in Education and Psychology*. New York: Reinhart and Winston, Inc.
- Nitko, A. J., Brookkhart, S. M. (1996). *Educational Assessment of Students*. New Jersey: Pearson Merrill Prentice Hall.
- Sax, G. (1980). *Principles of Educational and Psychological Measurement and Evaluation*. San francisco: Phoenix Publishing services.
- Suryabrata, Sumadi. (2005). *Pengembangan Alat Ukur Psikologis*. Yogyakarta: Andi Offset.
- Wright, B.D. & Stone, M.H (1979). *Best Test Design*. Chicago: Mes A Press.
- Wiberg, M. (2004). An Evaluation of Theory Test in The Swedish Driving-licence Test. *Classical Test Theory vs Item Response Theory*. Diambil tanggal 1 September 2009, dari <http://www8.umu.se/edmeas/publikationer/pdf/>
- _____. (2003). Undang-Undang RI Nomor 20, tahun 2003, tentang Sistem Pendidikan Nasional.
- _____. (2005). Peraturan Pemerintah RI Nomor 19, Tahun 2005, tentang Standar Nasional Pendidikan.
- _____. (2007). Peraturan Menteri RI Nomor 20 tahun 2007, Tentang Standar Penilaian Pendidikan.